International Annals of Science
ISSN: 2456-7132
Volume 8, Issue 1, pp. 1-11, 2020
DOI: https://doi.org/10.21467/ias.8.1.1-11

RESEARCH ARTICLE

Identification of Novel Key Biomarkers in Simpson-Golabi-Behmel Syndrome (SGBS): Evidence from Bioinformatics Analysis

Mujahed I. Mustafa[1*], Abdelrahman H. Abdelmoneim[1], Nafisa M. Elfadol[1], Naseem S. Murshed[1], Zainab O. Mohammed[2], Mohamed A. Hassan[1]

[1] Department of Biotechnology, Africa City of Technology, Sudan.
[2] Hematology Department, Ribat University Hospital, Sudan.

* Corresponding author email: mujahedimustafa@gmail.com

Received: 26 May 2019 / Revised: 10 July 2019 / Accepted: 12 July 2019 / Published: 21 July 2019

ABSTRACT

The Simpson-Golabi-Behmel Syndrome (SGBS) or overgrowth Syndrome is an uncommon genetic X-linked disorder highlighted by macrosomia, renal defects, cardiac weaknesses and skeletal abnormalities. The purpose of the work was to classify the functional nsSNPs of GPC3 to serve as genetic biomarkers for overgrowth syndrome. The raw data of *GPC3* gene were retrieved from dbSNP database and used to examine the most damaging effect using eight functional analysis tools, while we used I-mutant and MUPro to examine the effect of SNPs on GPC3 protein structure; The 3D structure of GPC3 protein is not found in the PDB, so RaptorX was used to create a 3D structural prototype to visualize the amino acids alterations by UCSF Chimera; For biophysical validation we used project HOPE; Lastly we run conservational analysis by BioEdit and Consurf web server respectively. Our results revealed three novel missense mutations (rs1460413167, rs1295603457 and rs757475450) that are that are more likely to be responsible for disturbance in the function and structure of GPC3. This work provides new insight into the molecular basis of overgrowth Syndrome by evidence from bioinformatics analysis. Three novel missense mutations (rs757475450, rs1295603457 and rs1460413167) are more likely to be responsible for disturbance in the function and structure of GPC3; therefore, they may be assisting as genetic biomarkers for overgrowth syndrome. As well as these SNPs can be used for the larger population-based studies of overgrowth syndrome.

Keywords: Bioinformatics analysis; Diagnostic markers; *GPC;* nsSNPs; Overgrowth syndrome.

1 Introduction

The Simpson-Golabi-Behmel Syndrome (SGBS) or Overgrowth Syndrome is an uncommon genetic disorder characterized by macrosomia, renal defects, cardiac weaknesses and skeletal abnormalities.[1-4] the first case was reported around 1940.[5] So far, two unlike types of overgrowth Syndrome have been defined. The typical SBGS type one [2-8] and a fatal and rare system, possibly 10 conditions defined known as SGBS type two.[9-11] Furthermore, these cases could rapidly develop Wilms' malignancy.[12] Early passing is more common.[13] Different mutations have been described in SGBS type one.[14-22]

Overgrowth Syndrome caused by mutations in glypican 3 (*GPC3)* gene is localized on Xq26.1 [23, 24] which encrypts glypican-3. [17, 19, 20, 25-29] that seemingly acting a bad part in growth control by an anonymous fate, However, outcomes from an exhaustive qualified study of growth forms in dual mutants missing GPC3 provided conclusive genetic evidence inconsistent with the theory that GPC3 performances as a growth suppressor.[29] Such a proteoglycan is contingent to show a vital part in regulate and diagnosis in mesodermal tissues and in tumors predisposition.[30, 31] Some studies show association between *GPC3* gene and some types of human cancers.[32-36]

The aim of this work was to detect the most deleterious SNPs in *GPC3* that may cause overgrowth Syndrome type one by using of different bioinformatics tools, Furthermore and to be used as genetic biomarkers. nsSNP is an alteration that occurs in a single base pair of amino acid which leads to disturb or change the corresponding protein's function, if the second possibility happened, it may cause a severe phenotypic impact and in return responsible for the pathology of the disease [37, 38] Clinical testing for deleterious SNPs frequently discloses alterations that are not easily considered as deleterious, for that reason a great effort has been done by translational bioinformatics tools for analysis of nsSNPs which have improved significantly in recent years and thus become more reliable for SNPs analysis.[39] Translational analysis has been considered as an essential science in the field of personalized medicine which aims to fill the gap between clinical and academic research by prioritizing the most pathogenic nsSNPs for further studies.[40-44] This is the first computational analysis of *GPC3* gene that classify nsSNPs for the larger population-based studies of overgrowth syndrome.

2 Methods

2.1 Data Mining

The raw data of *GPC3* gene were retrieved from National Center for Biotechnology Information (NCBI) website.[45, 46] The reference sequence of the protein was retrieved from Uniprot database.[47]

2.2 Functional Analysis

2.2.1 SIFT

It is the first functional analysis online tool which designed to predict whether a SNP is damaging or not by specific algorithm have a score <0.05 are predicted to be damaging SNP, otherwise it reflected to be not damaging.[48, 49]

2.2.2 PolyPhen

It is a functional analysis online tool to examine potential influences of a SNP on functional and structural characteristics of our protein of interest.[50, 51]

2.2.3 PROVEAN

It is a functional analysis online tool which we used to calculate if a SNP has an impression on the physical role of our protein of interest. PROVEAN probability has two possibilities, deleterious or neutral with cutoff -2.5.[52]

2.2.4 SNAP

It is a functional analysis tool with an artificial intelligence machine device called "neural network"; It distinguishes between effect and neutral variants/non-synonymous SNPs by taking a variety of sequence and variant features into account. [53, 54]

2.2.5 SNPs&GO

It is a functional analysis tool which distinguishes between the damaging SNPs from the neutral ones. The other methods were used too (PHD-SNP and PANTHER).[55, 56]

2.2.6 P-Mut

It is an online functional analysis tool for the clarification of amino acid alternates on proteins, permits the swift and accurate intention (80%) of the obsessive characteristics of each SNP stranded on the preparation of neural systems.[57, 58]

2.3 Stability Analysis

2.3.1 I-Mutant 3.0

It is SVM-based (Support Vector Machine) tool for the automatic prediction of protein stability changes upon single point mutations. The predictions are performed starting either from the protein structure or, more importantly, from the protein sequence.[59, 60]

2.3.2 MUPro

It is an online tool we used; it runs by the same concept of I-Mutant 3.0 but it's more accurate than I-Mutant 3.0 by 84.2%.[61, 62]

2.4 Biophysical and Visualization Analysis

2.4.1 Project Hope

It is an online web-server for biophysical validation which brings together a series of related protein data to form a model if there are enough 3D structural data; also to run this data to predict if the amino acid alteration may affect in the protein function or not.[63]

2.4.2 RaptorX

The 3D structure of the protein of GPC3 it is not found at protein data bank (PDB), so RaptorX was used to perform a 3D structure model for GPC3 protein.[64, 65]

2.4.3 UCSF Chimera

It is a visualization analysis program of 3D structure model, docking analysis and so many related analyses; the predicted model was used to visualize and compare the amino acid alterations by UCSF Chimera [66, 67].

2.5 Conservational Analysis

2.5.1 BioEdit

It is a program package created to stream a distinct program that can run approximately any sequences operation, demonstrating, as well as a few basic alignment studies.[68]

2.5.2 ConSurf Server

It is proposing evolutionary conservation outlines for proteins of known structure in the PDB. ConSurf red flag the similar amino acid sequences and run multi alignment approaches. The conserved regions of amino acids identify its site by using particular system.[69, 70]

3 Results

The effect of each SNP has been studied regarding to function and stability of the protein by different computational analysis tools with different considerations and features, in order to decrease the error to the lowest ratio possible (Figure 1).

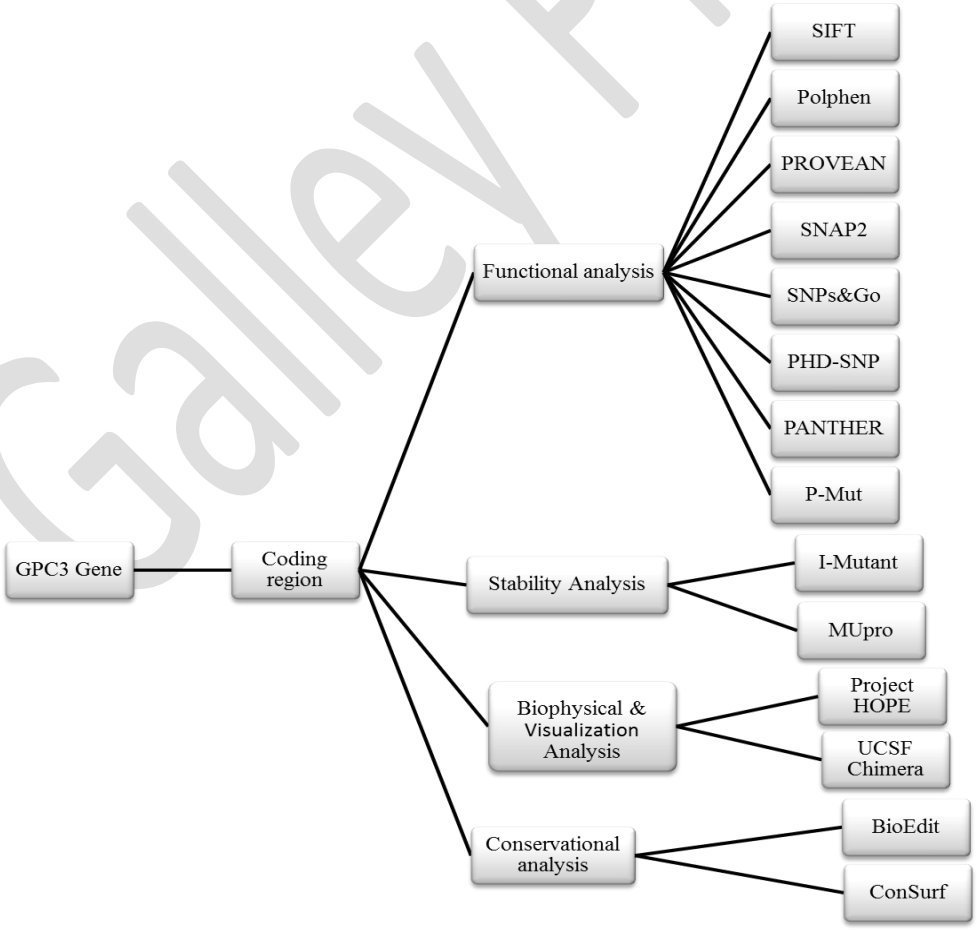

Figure 1: *Illustrative Workflow used for SNPs analysis*

Table 1: *Affecting protein function mutations predicted by several online tools:*

dbSNP rs#	sub	SIFT Prediction	Score	Polyphen Prediction	Score	PROVEAN Prediction	Score	SNAP2 Prediction	Score
-	D500Y	Deleterious	0	Damaging	1	Deleterious	-4.572	Deleterious	68
rs1203009272	G440R	Deleterious	0	Damaging	1	Deleterious	-5.147	Deleterious	90
rs104894854	W296R	Deleterious	0	Damaging	1	Deleterious	-13.836	Deleterious	98
rs1460413167	P212H	Deleterious	0	Damaging	1	Deleterious	-7.77	Deleterious	73
rs140848049	F208L	Deleterious	0	Damaging	1	Deleterious	-5.913	Deleterious	78
rs1295603457	C65Y	Deleterious	0	Damaging	1	Deleterious	-7.572	Deleterious	87
rs757475450	R39C	Deleterious	0	Damaging	1	Deleterious	-4.575	Deleterious	59

***Sub:** Substitutions

Table 2: *Disease related nsSNPs predicted by several online tools*

sub	SNPs&GO Prediction	RI	Probability	PANTHER Prediction	RI	Probability	PHD-SNP Prediction	RI	Probability	P-mut Prediction	Probability
W296R	Disease	7	0.858	Disease	10	0.99	Disease	9	0.925	Disease	0.89 (92%)
P212H	Disease	4	0.709	Disease	10	0.994	Disease	6	0.788	Disease	0.81 (89%)
C65Y	Disease	6	0.787	Disease	8	0.878	Disease	7	0.85	Disease	0.80 (89%)
R39C	Disease	3	0.63	Disease	8	0.903	Disease	4	0.687	Disease	0.73 (87%)

***RI:** Reliability Index

Table 3: *Structural investigation expected by I-mutant and MUPro:*

dbSNP rs#	Substitutions	SVM Prediction Effect	RI	Prediction	MUPro Prediction	Score
rs104894854	W296R	Decrease	8	-0.99	Decrease	-0.95856
rs1460413167	P212H	Decrease	8	-1.67	Decrease	-1.30249
rs1295603457	C65Y	Increase	0	-0.09	Decrease	-0.56309
rs757475450	R39C	Decrease	3	-0.74	Decrease	-0.28424

The total number of SNPs regarding to *GPC3* gene is 765 SNPs, out of 256 nsSNPs were submitted to SIFT, PolyPhen-2, PROVEAN and SNAP2 respectively. SIFT predicted 109 damaging mutations, PolyPhen-2 predicted 115 deleterious mutations (50 possibly damaging (less confident prediction) and 65 probably damaging (more confident prediction)), PROVEAN predicted 82 deleterious mutations and SNAP2 predicted 127 damaging mutations. Once we filtered the four positive deleterious mutations, the number of SNPs reduced to 7. (Table 1) after that, the same 7 mutations were submitted to SNPs&GO, PHD-SNP, PANTHER and P-Mut for further study to examine their influence on the function of *GPC3*; 7 deleterious mutations were predicted by PHD-SNP and P-mut, SNP&GO predicted 5, while PANTHER predicted 5 deleterious mutations. Once we filtered the four positive deleterious mutations the number reduced to 4 SNPs (Table 2) after that, we submitted them to I-Mutant and MUPro to investigate their effect on the stability; The two online tools revealed that, All the mutations decreased the protein stability, except for one SNP (G257D) was predicted by I-Mutant to increase the stability of the protein (Table 3).

4 Discussion

A significant interest in *Homo sapiens* genome has been focused to classify the deleterious SNPs; those are more likely to be responsible for inherited disorders. Therefore, a good effort was dictated to identify the most deleterious SNPs that may cause overgrowth syndrome. Our analysis revealed three novel SNPs in *GPC3*gene which were classified as highly deleterious SNPs, which as crucial impact at the functional level of the *GPC3* gene, our analysis based on different sequence and structure-based algorithms, Figure (1).

There is a study that has been reported which shows a missense mutation that causes overgrowth syndrome; [19] which matches with

this study findings. Some studies show association between *GPC3* gene and some types of liver cancer such as hepatocellular carcinoma.[30, 34, 71, 72] Therefore, this study can open the door for novel diagnostic biomarkers for hepatocellular carcinoma. Combination detection of serum GPC3 and pathogenic SNPs through clinical and genetic testing must be positively matched; this can enhance accuracy and efficiency of hepatocellular carcinoma diagnosis. In addition it confirms that (W296R) is pathogenic; this result matches with the result found previously in dbSNPs database. Furthermore, these mutations (P212H, C65Y, R39C) were recovered as untested, in this study were found to be all pathogenic.

At the functional level analysis, our results showed that all these nsSNPs substitutions (D500Y, G440R, W296R, P212H, F208L, C65Y, and R39C) were classified as likely pathogenic mutations, Table (1) the prediction efficacy has been increased by integrating the results of SIFT, PolyPhen-2, PROVEAN and SNAP2 based approaches, by combining the predictions of SNPs&GO, PhD-SNP, PANTHER, and P-Mut, Table (2) the output showed that all these nsSNPs (W296R, P212H, C65Y and R39C) are classified as highly pathogenic mutations. Therefore, our functional analysis suggested that these four nsSNPs might disrupt both the protein function and structure; while at the structural level analysis, MUPro results showed a decrease in stability for All these SNPs (W296R, P212H, C65Y and R39C) while I-Mutant results showed a decrease in stability for these SNPs (W296R, P212H and R39C), thus suggesting that these mutations could directly or indirectly destabilize the amino acid interactions triggering functional deviations of protein to some point. Table (3)

The most four deleterious SNPs were submitted to project HOPE which shown that all they are located in a domain of GPC3 protein; therefore, they may have a dynamic alteration in the protein function; In (Figure 2): (R39C): Shows the schematic structures of the original amino acid (in the left) which is Arginine and the mutant one (in the right) which is Cysteine. The backbone, which is the same for each amino acid, is colored red (in the green and red boxes) and the side chain, unique for each amino acid, is colored black. In addition, figure shows Close-up angle of the mutation. The protein is colored white, wide type residue colored green and mutant one colored red in position 39. The mutant residue is smaller than the wild-type residue; the wild-type residue charge was positive, while the mutant residue charge is neutral, this can cause loss of interactions with other molecules or residues; the mutant residue is more hydrophobic than the wild-type residue, and this can result in loss of hydrogen bonds and/or disturb correct folding.

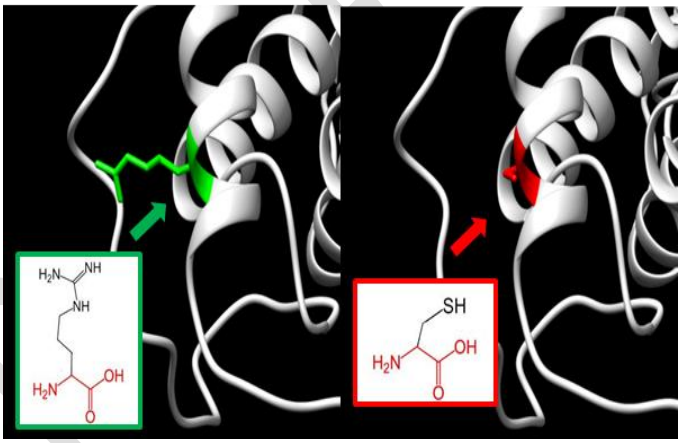

Figure 2: *(rs757475450) (R39C) Arginine changes to Cysteine at position 39; illustrated by chimera (v 1.8) and project HOPE.*

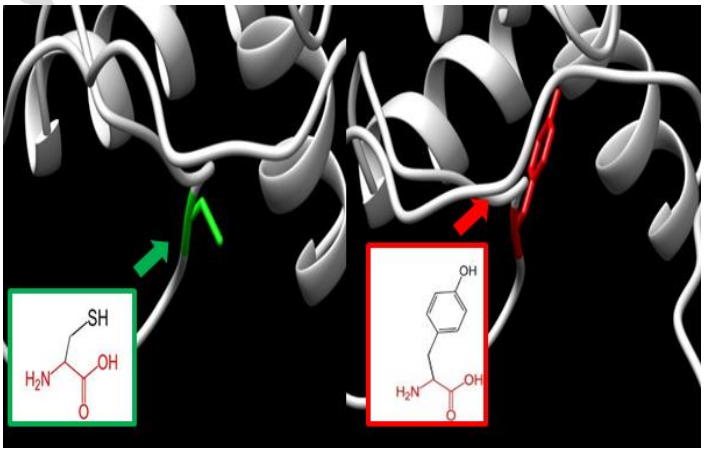

Figure 3: *(rs1295603457): (C65Y) Cysteine changes to Tyrosine at position 65; illustrated by chimera (v 1.8) and project HOPE.*

In (Figure 3): (C65Y): Shows the schematic structures of the original amino acid (in the left) which is Cysteine and the mutant one (in the right) which is Tyrosine. The backbone, which is the same for each amino acid, is colored red (in

the green and red boxes) and the side chain, unique for each amino acid, is colored black. In addition, figure shows Close-up angle of the mutation. The protein is colored white, wide type residue colored green and mutant one colored red in position 65. The wild-type and mutant amino acids differ in size; the mutant residue is bigger, this might lead to bumps; the hydrophobicity of the wild-type and mutant residue differs; hydrophobic interactions, either in the core of the protein or on the surface, will be lost.

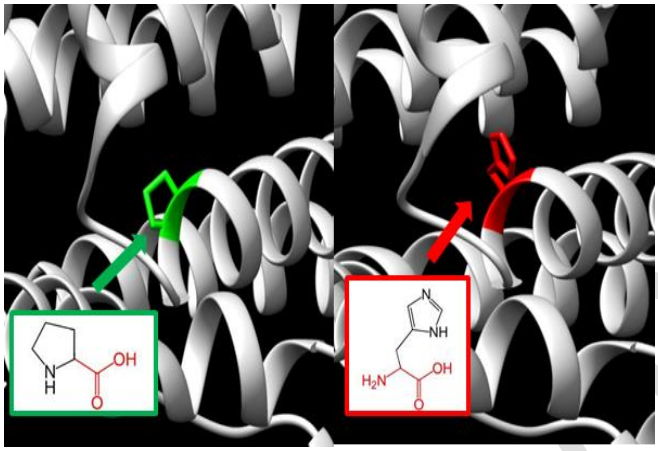

Figure 4: *(rs1460413167): (P212H) Proline changes to Histidine at position 212; illustrated by chimera (v 1.8) and project HOPE.*

In (Figure 4): (P212H): Shows the schematic structures of the original amino acid (in the left) which is Proline and the mutant one (in the right) which is Histidine. The backbone, which is the same for each amino acid, is colored red (in the green and red boxes) and the side chain, unique for each amino acid, is colored black. In addition, figure shows Close-up angle of the mutation. The protein is colored white, wide type residue colored green and mutant one colored red in position 212. The mutant residue is bigger, this might lead to bumps. The hydrophobicity of the wild-type and mutant residue differs; hydrophobic interactions, either in the core of the protein or on the surface, will be lost. Prolines are known to have a very rigid structure, sometimes forcing the backbone in a specific conformation. Possibly, this mutation changes a proline with such a function into another residue (Histidine), thereby disturbing the structure.

In (Figure 5): (W296R): Shows the schematic structures of the original amino acid (in the left) which is Tryptophan and the mutant one (in the right) which is Arginine. The backbone, which is the same for each amino acid, is colored red (in the green and red boxes) and the side chain, unique for each amino acid, is colored black. In addition, figure shows Close-up angle of the mutation. The protein is colored white, wide type residue colored green and mutant one colored red in position 296.

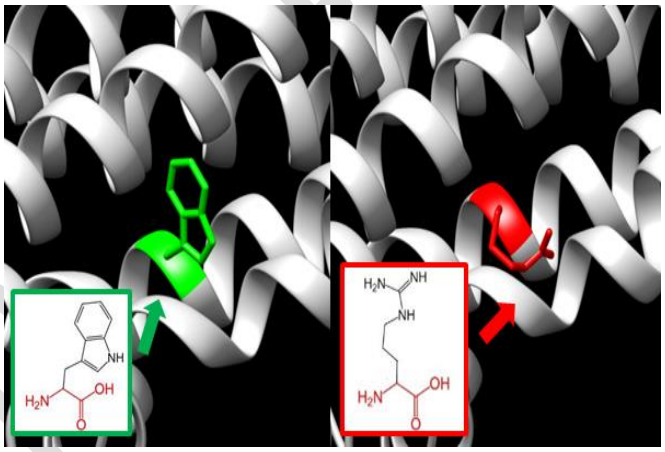

Figure 5: *(rs104894854): (W296R) Tryptophan changes to Arginine at position 296; illustrated by chimera (v1.8) and project HOPE.*

The wild-type residue charge was neutral, the mutant residue charge is positive, the mutation introduces a charge, and this can cause repulsion of ligands or other residues with the same charge; the wild-type and mutant amino acids differ in size, and the mutant residue is smaller, this might lead to loss of interactions; The hydrophobicity of the wild-type and mutant residue differs, hydrophobic interactions, either in the core of the protein or on the surface, will be lost.

We also observed that, all the four SNPs were located in conserve region. We believe that amino acids conserved across species are playing a crucial role at the functional level; therefore, the four SNPs that we have detected are more probable disease causing ones; (Figure 6) The same results were confirmed by ConSurf, which show the nsSNPs that they are located at extremely conserved sites; therefore, we have confidence that these SNPs have a tendency to

be the most deleterious SNPs that may cause overgrowth syndrome. (Figure 7)

This study is the first computational approach while all other earlier studies were in vitro, in vivo and whole exome sequencing. [73-76] It revealed three novel missense mutations that are more likely to be responsible for disturbance in the function and structure of GPC3; therefore, they could be used as diagnostic markers to Predict overgrowth syndrome.[77] Lastly, some appreciations of wet lab techniques are suggested to support our in silico analysis findings.

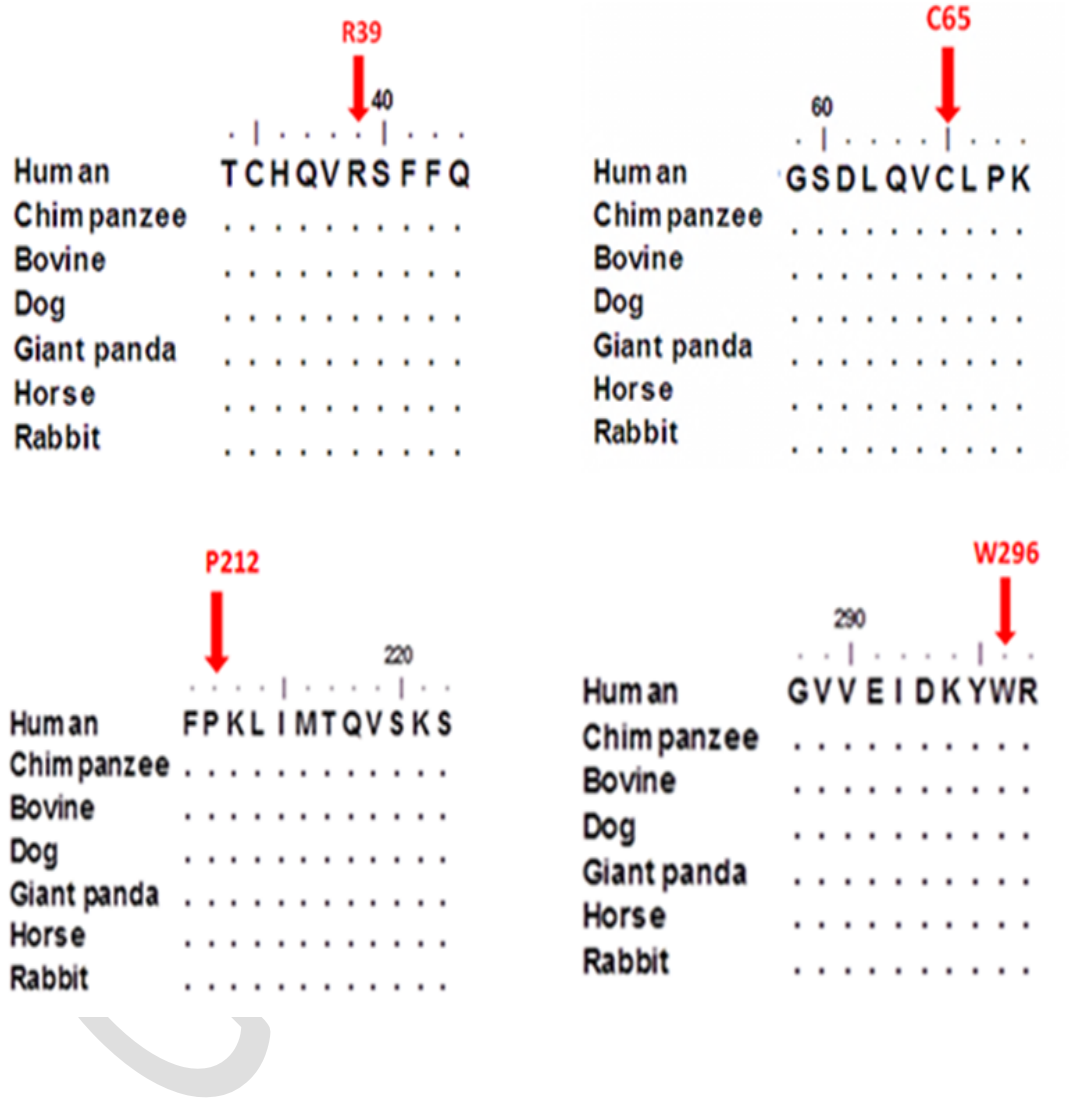

Figure 6: *GPC3 Family of seven protein sequences representing that, the normal amino acids are expected to be altered (showed by red arrows) are evolutionarily conserved across species.*

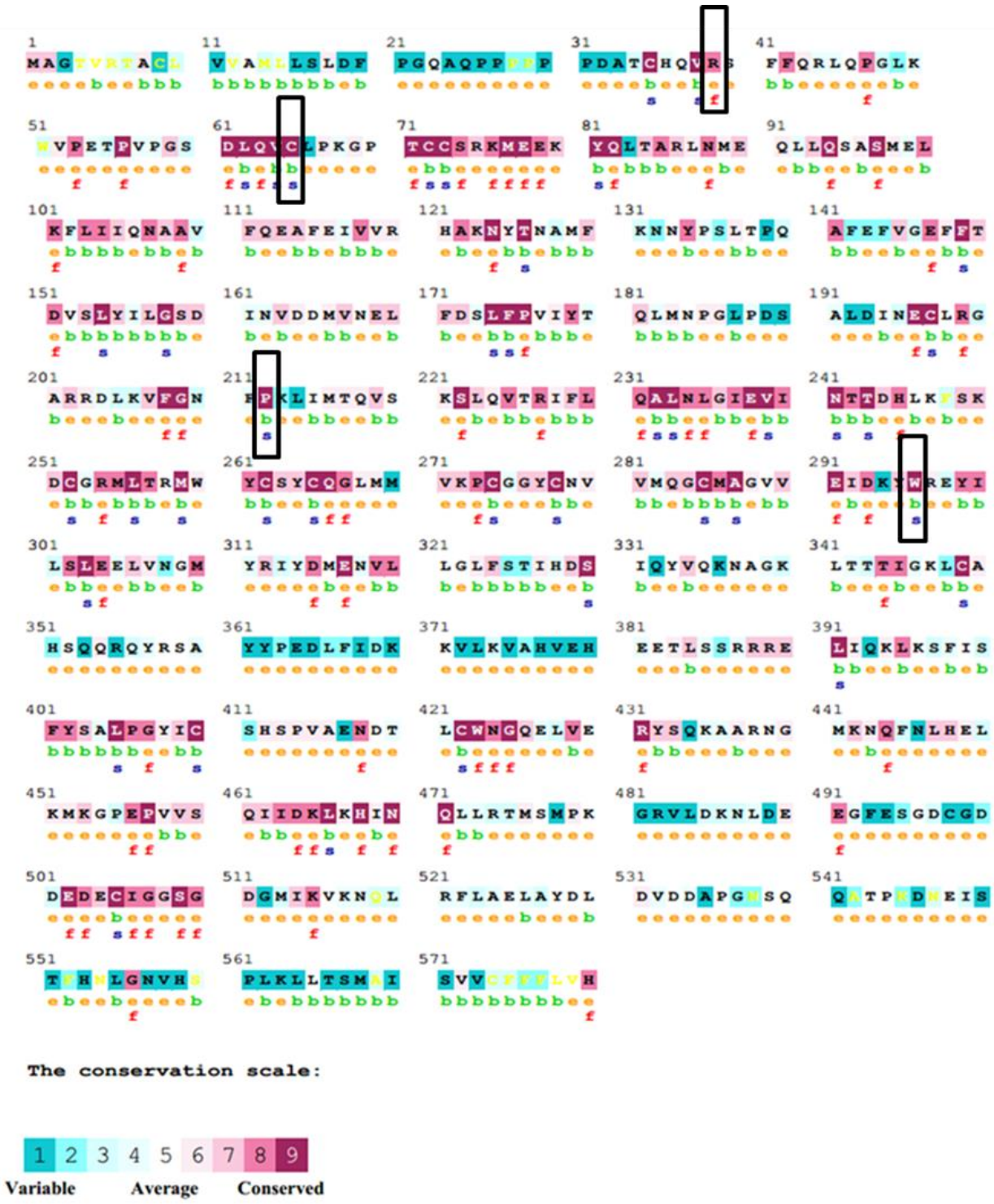

Figure 7: *Shows the conserved amino acids across species in GPC3 protein were determined using Consurf. (e) An exposed residue according to the neural-network algorithm via an orange letter. (b) Residues predicted to be buried are demonstrated via a green letter. (f) A predicted functional residue (highly conserved and exposed) are indicated with a red letter. (s) A predicted structural residue (highly conserved and buried) that are demonstrated with a blue letter. (I) Insufficient data- the calculation for this site was performed on less than 10% of the sequences are demonstrated via a yellow letter.*

5 Conclusion

Functional and structural impact of SNPs in the *GPC3* gene was found out by using computational prediction tools; Out of a total of 765 SNPs in the *GPC3* gene, 256 were nsSNPs; out of 256 missense nsSNPs, 4 were found to be the most deleterious nsSNPs (three of them were novel R39C (rs757475450), C65Y (rs1295603457), and P212H (rs1460413167)) by eight functional analysis tools. Stability analysis results showed that the amino acid residue substitutions which had the greatest impact on the stability of the GPC3 protein were mutations R39C (rs757475450), C65Y (rs1295603457), P212H (rs1460413167) and W296R (rs104894854). This result helped us to characterize the impact of nsSNPs on *GPC3* gene and should be considered important candidates in causing of overgrowth syndrome.

6 Declarations

6.1 Acknowledgment

The authors wish to acknowledge the enthusiastic cooperation of Africa City of Technology - Sudan.

6.2 Competing Interests

The authors declare that no conflict of interest exist in this publication.

7 How to Cite this Article:

Will be updated in the final version.

References

[1] E. Cottereau, I. Mortemousque, M. P. Moizard, L. Burglen, D. Lacombe, B. Gilbert-Dussardier, *et al.*, "Phenotypic spectrum of Simpson-Golabi-Behmel syndrome in a series of 42 cases with a mutation in GPC3 and review of the literature," *Am J Med Genet C Semin Med Genet*, vol. 163c, pp. 92-105, May 2013.

[2] A. Behmel, E. Plochl, and W. Rosenkranz, "A new X-linked dysplasia gigantism syndrome: identical with the Simpson dysplasia syndrome?," *Hum Genet*, vol. 67, pp. 409-13, 1984.

[3] M. Golabi and L. Rosen, "A new X-linked mental retardation-overgrowth syndrome," *Am J Med Genet*, vol. 17, pp. 345-58, Jan 1984.

[4] M. R. DeBaun, J. Ess, and S. Saunders, "Simpson Golabi Behmel syndrome: progress toward understanding the molecular basis for overgrowth, malformation, and cancer predisposition," *Mol Genet Metab*, vol. 72, pp. 279-86, Apr 2001.

[5] F. Gurrieri, M. G. Pomponi, R. Pietrobono, E. Lucci-Cordisco, E. Silvestri, G. Storniello, *et al.*, "The Simpson-Golabi-Behmel syndrome: A clinical case and a detective story," *Am J Med Genet A*, vol. 155a, pp. 145-8, Jan 2011.

[6] J. L. Simpson, S. Landey, M. New, and J. German, "A previously unrecognized X-linked syndrome of dysmorphia," *Birth Defects Orig Artic Ser*, vol. 11, pp. 18-24, 1975.

[7] M. L. Vuillaume, M. P. Moizard, S. Rossignol, E. Cottereau, S. Vonwill, J. L. Alessandri, *et al.*, "Mutation update for the GPC3 gene involved in Simpson-Golabi-Behmel syndrome and review of the literature," *Hum Mutat*, vol. 39, pp. 2110-2112, Dec 2018.

[8] G. Neri and M. Moscarda, "Overgrowth syndromes: a classification," *Endocr Dev*, vol. 14, pp. 53-60, 2009.

[9] J. Tenorio, P. Arias, V. Martinez-Glez, F. Santos, S. Garcia-Minaur, J. Nevado, *et al.*, "Simpson-Golabi-Behmel syndrome types I and II," *Orphanet J Rare Dis*, vol. 9, p. 138, Sep 20 2014.

[10] L. M. Brzustowicz, S. Farrell, M. B. Khan, and R. Weksberg, "Mapping of a new SGBS locus to chromosome Xp22 in a family with a severe form of Simpson-Golabi-Behmel syndrome," *Am J Hum Genet*, vol. 65, pp. 779-83, Sep 1999.

[11] D. Terespolsky, S. A. Farrell, J. Siegel-Bartelt, and R. Weksberg, "Infantile lethal variant of Simpson-Golabi-Behmel syndrome associated with hydrops fetalis," *Am J Med Genet*, vol. 59, pp. 329-33, Nov 20 1995.

[12] R. M. Hughes-Benzie, A. G. Hunter, J. E. Allanson, and A. E. Mackenzie, "Simpson-Golabi-Behmel syndrome associated with renal dysplasia and embryonal tumor: localization of the gene to Xqcen-q21," *Am J Med Genet*, vol. 43, pp. 428-35, Apr 15-May 1 1992.

[13] C. L. Garganta and J. N. Bodurtha, "Report of another family with Simpson-Golabi-Behmel syndrome and a review of the literature," *Am J Med Genet*, vol. 44, pp. 129-35, Sep 15 1992.

[14] C. B. Griffith, R. C. Probert, and G. H. Vance, "Genital anomalies in three male siblings with Simpson-Golabi-Behmel syndrome," *Am J Med Genet A*, vol. 149a, pp. 2484-8, Nov 2009.

[15] A. Vaisfeld, M. G. Pomponi, R. Pietrobono, E. Tabolacci, and G. Neri, "Simpson-Golabi-Behmel syndrome in a female: A case report and an unsolved issue," *Am J Med Genet A*, vol. 173, pp. 285-288, Jan 2017.

[16] S. Halayem, M. Hamza, F. Maazoul, H. Ben Turkia, M. Touati, N. Tebib, *et al.*, "Distinctive findings in a boy with Simpson-Golabi-Behmel syndrome," *Am J Med Genet A*, vol. 170a, pp. 1035-9, Apr 2016.

[17] M. Li, C. Shuman, Y. L. Fei, E. Cutiongco, H. A. Bender, C. Stevens, *et al.*, "GPC3 mutation analysis in a spectrum of patients with overgrowth expands the phenotype of Simpson-Golabi-Behmel syndrome," *Am J Med Genet*, vol. 102, pp. 161-8, Aug 1 2001.

[18] S. Mariani, L. Iughetti, R. Bertorelli, D. Coviello, M. Pellegrini, A. Forabosco, *et al.*, "Genotype/phenotype correlations of males affected by Simpson-Golabi-Behmel syndrome with GPC3 gene mutations: patient report and review of the literature," *J Pediatr Endocrinol Metab*, vol. 16, pp. 225-32, Feb 2003.

[19] M. Veugelers, B. D. Cat, S. Y. Muyldermans, G. Reekmans, N. Delande, S. Frints, *et al.*, "Mutational

analysis of the GPC3/GPC4 glypican gene cluster on Xq26 in patients with Simpson-Golabi-Behmel syndrome: identification of loss-of-function mutations in the GPC3 gene," *Hum Mol Genet*, vol. 9, pp. 1321-8, May 22 2000.

[20] G. Pilia, R. M. Hughes-Benzie, A. MacKenzie, P. Baybayan, E. Y. Chen, R. Huber, *et al.*, "Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome," *Nat Genet*, vol. 12, pp. 241-7, Mar 1996.

[21] K. Ridnoi, E. Kurvinen, S. Pajusalu, T. Reimand, and K. Ounap, "Two Consecutive Pregnancies with Simpson-Golabi-Behmel Syndrome Type 1: Case Report and Review of Published Prenatal Cases," *Mol Syndromol*, vol. 9, pp. 205-213, Jul 2018.

[22] H. K. Stove, N. Becher, V. Gjorup, M. Ramsing, I. Vogel, and E. M. Vestergaard, "First reported case of Simpson-Golabi-Behmel syndrome in a female fetus diagnosed prenatally with chromosomal microarray," *Clin Case Rep*, vol. 5, pp. 608-612, May 2017.

[23] G. Jedraszak, M. Girard, A. Mellos, D. D. Djeddi, C. Chardot, A. Vanrenterghem, *et al.*, "A patient with Simpson-Golabi-Behmel syndrome, biliary cirrhosis and successful liver transplantation," *Am J Med Genet A*, vol. 164a, pp. 774-7, Mar 2014.

[24] R. Kosaki, T. Takenouchi, N. Takeda, M. Kagami, K. Nakabayashi, K. Hata, *et al.*, "Somatic CTNNB1 mutation in hepatoblastoma from a patient with Simpson-Golabi-Behmel syndrome and germline GPC3 mutation," *Am J Med Genet A*, vol. 164a, pp. 993-7, Apr 2014.

[25] R. Huber, L. Crisponi, R. Mazzarella, C. N. Chen, Y. Su, H. Shizuya, *et al.*, "Analysis of exon/intron structure and 400 kb of genomic sequence surrounding the 5'-promoter and 3'-terminal ends of the human glypican 3 (GPC3) gene," *Genomics*, vol. 45, pp. 48-58, Oct 1 1997.

[26] J. Schmidt, R. Hollstein, F. J. Kaiser, and G. Gillessen-Kaesbach, "Molecular analysis of a novel intragenic deletion in GPC3 in three cousins with Simpson-Golabi-Behmel syndrome," *Am J Med Genet A*, vol. 173, pp. 1400-1405, May 2017.

[27] D. D. Villarreal, H. Villarreal, A. M. Paez, D. Peppas, J. Lynch, E. Roeder, *et al.*, "A patient with a unique frameshift mutation in GPC3, causing Simpson-Golabi-Behmel syndrome, presenting with craniosynostosis, penoscrotal hypospadias, and a large prostatic utricle," *Am J Med Genet A*, vol. 161a, pp. 3121-5, Dec 2013.

[28] S. Yano, B. Baskin, A. Bagheri, Y. Watanabe, K. Moseley, A. Nishimura, *et al.*, "Familial Simpson-Golabi-Behmel syndrome: studies of X-chromosome inactivation and clinical phenotypes in two female individuals with GPC3 mutations," *Clin Genet*, vol. 80, pp. 466-71, Nov 2011.

[29] E. Chiao, P. Fisher, L. Crisponi, M. Deiana, I. Dragatsis, D. Schlessinger, *et al.*, "Overgrowth of a mouse model of the Simpson-Golabi-Behmel syndrome is independent of IGF signaling," *Dev Biol*, vol. 243, pp. 185-206, Mar 1 2002.

[30] P. Lapunzina, I. Badia, C. Galoppo, E. De Matteo, P. Silberman, A. Tello, *et al.*, "A patient with Simpson-Golabi-Behmel syndrome and hepatocellular carcinoma," *J Med Genet*, vol. 35, pp. 153-6, Feb 1998.

[31] R. Savarirayan and A. Bankier, "Simpson-Golabi-Behmel syndrome and attention deficit hyperactivity disorder in two brothers," *J Med Genet*, vol. 36, pp. 574-6, Jul 1999.

[32] W. X. Bai, J. Gao, C. Qian, and X. Q. Zhang, "[A bioinformatics analysis of differentially expressed genes associated with liver cancer]," *Zhonghua Gan Zang Bing Za Zhi*, vol. 25, pp. 435-439, Jun 20 2017.

[33] F. Cartier, E. Indersie, S. Lesjean, J. Charpentier, K. B. Hooks, A. Ghousein, *et al.*, "New tumor suppressor microRNAs target glypican-3 in human liver cancer," *Oncotarget*, vol. 8, pp. 41211-41226, Jun 20 2017.

[34] Y. Wu, H. Liu, and H. Ding, "GPC-3 in hepatocellular carcinoma: current perspectives," *J Hepatocell Carcinoma*, vol. 3, pp. 63-67, 2016.

[35] J. Filmus, "Glypicans in growth control and cancer," *Glycobiology*, vol. 11, pp. 19r-23r, Mar 2001.

[36] H. Lin, R. Huber, D. Schlessinger, and P. J. Morin, "Frequent silencing of the GPC3 gene in ovarian cancer cell lines," *Cancer Res*, vol. 59, pp. 807-10, Feb 15 1999.

[37] G. Shaw, "Polymorphism and single nucleotide polymorphisms (SNPs)," *BJU Int*, vol. 112, pp. 664-5, Sep 2013.

[38] Y. Chen, F. Cunningham, D. Rios, W. M. McLaren, J. Smith, B. Pritchard, *et al.*, "Ensembl variation resources," *BMC Genomics*, vol. 11, p. 293, May 11 2010.

[39] C. George Priya Doss, C. Sudandiradoss, R. Rajasekaran, P. Choudhury, P. Sinha, P. Hota, *et al.*, "Applications of computational algorithm tools to identify functional SNPs," *Funct Integr Genomics*, vol. 8, pp. 309-16, Nov 2008.

[40] J. D. Tenenbaum, "Translational Bioinformatics: Past, Present, and Future," *Genomics Proteomics Bioinformatics*, vol. 14, pp. 31-41, Feb 2016.

[41] J. Vamathevan and E. Birney, "A Review of Recent Advances in Translational Bioinformatics: Bridges from Biology to Medicine," *Yearb Med Inform*, vol. 26, pp. 178-187, Aug 2017.

[42] P. Katara, "Single nucleotide polymorphism and its dynamics for pharmacogenomics," *Interdiscip Sci*, vol. 6, pp. 85-92, Jun 2014.

[43] J. Wang, G. S. Pang, S. S. Chong, and C. G. Lee, "SNP web resources and their potential applications in personalized medicine," *Curr Drug Metab*, vol. 13, pp. 978-90, Sep 1 2012.

[44] D. Gefel, I. Maslovsky, and J. Hillel, "[Application of single nucleotide polymorphisms (SNPs) for the detection of genes involved in the control of complex diseases]," *Harefuah*, vol. 147, pp. 449-54, 476, May 2008.

[45] D. A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, *et al.*, "GenBank," *Nucleic Acids Res*, vol. 45, pp. D37-d42, Jan 4 2017.

[46] "NCBI website."

[47] "UniProt: the universal protein knowledgebase," *Nucleic Acids Res*, vol. 45, pp. D158-d169, Jan 4 2017.

[48] N. L. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider, and P. C. Ng, "SIFT web server: predicting effects of amino acid substitutions on proteins," *Nucleic Acids Res*, vol. 40, pp. W452-7, Jul 2012.

[49] *SIFT server*. Available: https://sift.bii.a-star.edu.sg/

[50] E. Capriotti and R. B. Altman, "Improving the prediction of disease-related variants using protein three-dimensional structure," *BMC Bioinformatics*, vol. 12 Suppl 4, p. S3, 2011.

[51] *PolyPhen-2 server*. Available: http://genetics.bwh.harvard.edu/pph2/

[52] Y. Choi, G. E. Sims, S. Murphy, J. R. Miller, and A. P. Chan, "Predicting the functional effect of amino acid

substitutions and indels," *PLoS One*, vol. 7, p. e46688, 2012.

[53] M. Hecht, Y. Bromberg, and B. Rost, "Better prediction of functional effects for sequence variants," *BMC Genomics*, vol. 16 Suppl 8, p. S1, 2015.

[54] *SNAP2 server*. Available: https://rostlab.org/services/snap2web/

[55] R. Calabrese, E. Capriotti, P. Fariselli, P. L. Martelli, and R. Casadio, "Functional annotations improve the predictive score of human disease-related mutations in proteins," *Hum Mutat*, vol. 30, pp. 1237-44, Aug 2009.

[56] *SNPs&Go server*. Available: http://snps.biofold.org/snps-and-go/snps-and-go.html

[57] V. Lopez-Ferrando, A. Gazzo, X. de la Cruz, M. Orozco, and J. L. Gelpi, "PMut: a web-based tool for the annotation of pathological variants on proteins, 2017 update," *Nucleic Acids Res*, vol. 45, pp. W222-w228, Jul 3 2017.

[58] *P-Mut server*. Available: http://mmb.irbbarcelona.org/PMut

[59] E. Capriotti, P. Fariselli, and R. Casadio, "I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure," *Nucleic Acids Res*, vol. 33, pp. W306-10, Jul 1 2005.

[60] *I-Mutant server*. Available: http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi

[61] J. Cheng, A. Randall, and P. Baldi, "Prediction of protein stability changes for single-site mutations using support vector machines," *Proteins*, vol. 62, pp. 1125-32, Mar 1 2006.

[62] *MUPro server*. Available: http://mupro.proteomics.ics.uci.edu/

[63] *project HOPE server*. Available: http://www.cmbi.ru.nl/hope

[64] *RaptorX server*. Available: (http://raptorx.uchicago.edu/

[65] S. Wang, W. Li, S. Liu, and J. Xu, "RaptorX-Property: a web server for protein structure property prediction," *Nucleic Acids Res*, vol. 44, pp. W430-5, Jul 8 2016.

[66] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, *et al.*, "UCSF Chimera--a visualization system for exploratory research and analysis," *J Comput Chem*, vol. 25, pp. 1605-12, Oct 2004.

[67] *UCSF Chimera*. Available: http://www.cgl.ucsf.edu/chimera/

[68] *BioEdit*. Available: http://www.mbio.ncsu.edu/bioedit/bioedit.html

[69] H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, *et al.*, "ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules," *Nucleic Acids Res*, vol. 44, pp. W344-50, Jul 8 2016.

[70] *ConSurf server*. Available: http://consurf.tau.ac.il/).

[71] B. Sun, Z. Huang, B. Wang, Y. Yu, S. Lin, L. Luo, *et al.*, "Significance of Glypican-3 (GPC3) Expression in Hepatocellular Cancer Diagnosis," *Med Sci Monit*, vol. 23, pp. 850-855, Feb 16 2017.

[72] C. Chen, X. Huang, Z. Ying, D. Wu, Y. Yu, X. Wang, *et al.*, "Can glypican-3 be a disease-specific biomarker?," *Clin Transl Med*, vol. 6, p. 18, Dec 2017.

[73] P. Mochalski, E. Diem, K. Unterkofler, A. Mundlein, H. Drexel, C. A. Mayhew, *et al.*, "In vitro profiling of volatile organic compounds released by Simpson-Golabi-Behmel syndrome adipocytes," *J Chromatogr B Analyt Technol Biomed Life Sci*, vol. 1104, pp. 256-261, Jan 1 2019.

[74] A. X. Zhu, P. J. Gold, A. B. El-Khoueiry, T. A. Abrams, H. Morikawa, N. Ohishi, *et al.*, "First-in-man phase I study of GC33, a novel recombinant humanized antibody against glypican-3, in patients with advanced hepatocellular carcinoma," *Clin Cancer Res*, vol. 19, pp. 920-8, Feb 15 2013.

[75] A. Das Bhowmik and A. Dalal, "Whole exome sequencing identifies a novel frameshift mutation in GPC3 gene in a patient with overgrowth syndrome," *Gene*, vol. 572, pp. 303-6, Nov 10 2015.

[76] C. Kehrer, A. Hoischen, R. Menkhaus, E. Schwab, A. Muller, S. Kim, *et al.*, "Whole exome sequencing and array-based molecular karyotyping as aids to prenatal diagnosis in fetuses with suspected Simpson-Golabi-Behmel syndrome," *Prenat Diagn*, vol. 36, pp. 961-965, Oct 2016.

[77] Z. B. Alwi, "The Use of SNPs in Pharmacogenomics Studies," *Malays J Med Sci*, vol. 12, pp. 4-12, Jul 2005.